



LA VIRTUALISATION DES DONNÉES POUR LE

CDO

Les données ont toujours été un atout précieux, mais le directeur des données (Chief Data Officer - CDO) sait aujourd'hui que la valeur des données augmente si celles-ci peuvent être facilement combinées à d'autres ensembles de données. Le CDO sait aussi qu'avec l'émergence de nouveaux types de données - les données du cloud et du big data, les données de capteurs de l'Internet des objets (IoT), les données des flux de médias sociaux, ou même des données que l'on croyait inutilisables à des fins analytiques, comme les données transactionnelles et non structurées - leur valeur est potentiellement illimitée.

Cependant, le CDO est bien conscient des défis posés par l'intégration des données, en particulier entre les sources modernes et anciennes, et entre les données au repos et les données en mouvement. La virtualisation des données est une technologie moderne d'intégration des données qui résout la plupart, sinon la totalité, des problèmes d'intégration des données en s'appuyant sur une approche unique : au lieu de répliquer les données et de les déplacer vers un nouveau référentiel consolidé, la virtualisation des données connecte les utilisateurs et les applications à une vue des données, en temps réel, laissant les données source exactement là où elles se trouvent. En plus d'économiser sur les coûts de stockage, la virtualisation des données permet d'accéder à tout type de source de données, y compris celles qui sont traditionnellement incompatibles avec les entrepôts de données physiques.

Gartner a suivi avec attention l'évolution du paysage de l'intégration des données. Nous partageons ici les principales tendances en matière d'intégration de données publiées par Gartner, suivies d'une brève description de la façon dont les CDO, grâce à la virtualisation des données, peuvent s'inscrire dans cette tendance.

« D'ici 2021, 75 % des rapports standards seront remplacés ou complétés par des informations automatisées délivrées quand elles sont le plus nécessaires ». ¹

La virtualisation des données facilite la création de rapports sur une base ad hoc, ou selon les besoins, puisque la virtualisation des données ne nécessite pas la réplique physique des données pour créer un rapport intégré.



Seacoast Banking Corporation of Florida a tiré parti de la virtualisation des données pour combiner les données provenant de diverses sources et ainsi mettre en place une véritable infrastructure de BI (Business Intelligence) en self-service. Le nouveau système permet aux métiers de générer des rapports opérationnels et des rapports d'analyse d'entreprise beaucoup plus rapidement qu'auparavant. Ce qui aurait pris huit mois avec leur processus ETL (extraction, transformation et chargement) prend désormais cinq mois grâce à la virtualisation des données, soit un délai de mise sur le marché 40 % plus rapide.



« D'ici 2020, la plupart des cas d'usage data et analytiques nécessiteront une connexion à des sources de données distribuées, si bien que les entreprises devront doubler leurs investissements dans la gestion des métadonnées ».¹

La virtualisation des données agit comme une couche d'accès universelle, reliant les utilisateurs et les applications à une myriade de sources de données sous-jacentes. La couche de virtualisation des données elle-même ne contient aucune donnée, seulement les métadonnées nécessaires pour accéder aux différentes sources. En tant que telle, la virtualisation des données crée un accès transparent aux sources de données distribuées et une plate-forme unifiée pour gérer facilement les métadonnées afin de réduire les investissements dans la gestion des métadonnées.

« D'ici 2021, l'usage d'une sémantique de données agnostique basée sur de nouvelles technologies et de nouvelles pratiques réduira de 35 % les coûts de gestion et d'intégration des données ».¹

En établissant une couche d'accès universelle, la virtualisation des données facilite la création d'une sémantique conviviale qui masque les complexités d'accès pour l'utilisateur, par exemple, savoir où se trouvent les données ou disposer des informations d'identification de sécurité requises.



INDIANA UNIVERSITY

L'Université de l'Indiana a mis à profit la virtualisation des données pour soutenir un projet de gestion de l'information de grande envergure appelé Decision Support Initiative (DSI). L'objectif est de fournir des données pertinentes et exactes pour faciliter et améliorer la prise de décision au sein de l'Université. Parce que le projet DSI relie les parties prenantes aux sources de données distribuées, l'université s'est appuyée sur les capacités de gestion des métadonnées de la couche de virtualisation des données.



Vizient a tiré parti de la virtualisation des données pour établir un datamart financier virtuel qui permet d'obtenir des aperçus de données uniques et transparents sur une variété de sujets, réunis sous des dénominations courantes de type « Ventes des fournisseurs ». Le datamart dessert avec succès plus de 400 utilisateurs actifs répartis dans 6 départements.

« D'ici 2020, les entreprises ayant adopté des stratégies de hub de données pourront obtenir des résultats à partir de données partagées et régulées à un coût inférieur d'au moins 60 % ». ¹

Parce qu'elle sert de point d'accès universel pour toutes les sources de données d'entreprise, la virtualisation des données est la méthode la plus simple, la plus directe et la plus rentable pour mettre en œuvre un hub de données.

TransAlta

TransAlta a mis à profit la virtualisation des données pour implémenter un nouveau système de trading de l'énergie faisant office de hub de données. Celui-ci remonte les données de diverses sources internes et externes, dont certaines stockées dans le cloud et d'autres sur site auprès des organismes de réglementation. Le nouveau hub de données a permis de réduire la main-d'œuvre et les coûts inutiles et donné aux intervenants des capacités de self-service, réduisant ainsi leur dépendance vis-à-vis de l'IT.

« Entre 2016 et 2019, les dépenses d'analyse en temps réel augmenteront trois fois plus vite que les dépenses d'analyse en temps non réel ». ¹

Parce qu'elle permet d'accéder en temps réel à toutes les sources de données possibles, la virtualisation des données rend possible l'analyse en temps réel.

logitech

Logitech a tiré parti de la virtualisation des données pour intégrer en temps réel les sources de données dans le cloud et sur site, l'objectif étant d'alimenter des applications d'analyse et de reporting à partir d'une source de vérité unique. Le nouveau système permet aux utilisateurs de consommer l'information simplement, en self-service, et de manière agnostique vis-à-vis de l'outil, puisque tous les outils accèdent à la même vue unifiée des données. Grâce au prototypage rapide et à l'utilisation optimale des ressources, le nouveau système a permis de réduire considérablement les dépenses d'exploitation et les coûts élevés de réingénierie.

¹ Gartner: **100 Data and Analytics Predictions Through 2022.**